CHAPTER 3

RELIABILITY PREDICTION

3-1. Introduction to reliability prediction

It is unfortunate that the term "prediction" was ever used in connection with assessing the reliability of a design or product. Prediction has connotations of reading tea leaves or Tarot cards, or gazing into a crystal ball. Even if one compares reliability prediction to weather prediction, those unfamiliar with reliability but all too familiar with weather reports will form an uncomplimentary opinion of reliability prediction. A reliability prediction is nothing more than a quantitative assessment of the level of reliability inherent in a design or achieved in a test model or production product.

3-2. Uses of predictions

Although prediction is the subject of much controversy and debate, few people question the need to quantitatively assess the reliability of an item. Predictions are need for several reasons.

a. *Evaluate alternatives.* In creating a design, the engineer must decide on which parts, what materials, and, in coordination with the manufacturing/production engineers, the types of processes that will be used. Many factors influence these decisions, including costs, established lists of qualified parts and suppliers, existing manufacturing/production capabilities, and so forth. Reliability must also be a factor in selecting parts, materials, and processes. It is not necessary to always select the most reliable alternative. For example, it is not as important to use extremely reliable, and therefore expensive, parts, as it is to properly apply the parts that are selected. By using what is known as robust design techniques, even modestly reliable parts can be used in products where high reliability is required. Predictions assist in the process of evaluating alternatives.

b. *Provide a quantitative basis for design trade-offs.* In designing any product, but especially when designing complex systems such as those used by the military, it is seldom if ever possible to optimize all aspects of the product. It has been said that systems engineering is a process of compromises, in which individual performance parameters or characteristics may be sub-optimized to optimize the overall product performance. For example, a structure may need to be as light as possible but have extremely good fatigue characteristics and carry very high loads. These requirements conflict – maximizing any one may compromise another. Reliability is just one of many product requirements that must be considered in design trades. The most common trade is with the design characteristic of maintainability. That is, it may be possible to relax a reliability requirement if the time to repair can be decreased, thereby yielding the required level of system availability. Predictions help us make such trades on a quantitative basis.

c. *Compare established reliability requirements with state-of-the-art feasibility.* All too often, a requirement is levied on a supplier without determining if the requirement is realistic. Consequently, much time and resources are spent trying to achieve what is inherently unachievable. Although it is natural to <u>want</u> products and systems that are as reliable as possible, we must concentrate on the level of reliability that is <u>needed</u>, to stay within schedule and budget constraints. This level is the one that is dictated by mission and life cycle cost considerations, is achievable given the state of the art of the technology being used, and is consistent with the other system performance requirements. Predictions allow us to assess the feasibility of a requirement.

d. *Provide guidance in budget and schedule decisions.* Assessing the reliability of a design throughout the design process helps to determine if budgets and schedules are sufficient or, on the other hand, determine if we can achieve the required level of reliability within budget and schedule constraints. Early estimates of reliability can be important inputs into determining a program budget and schedule.

e. *Provide a uniform basis for proposal preparation, evaluation, and selection.* When multiple sources are available to bid on a new product or system contract, the customer must be able to select the best supplier. Obviously cost is one way of choosing between suppliers, provided all the suppliers can design and build a system with the required performance with the same level of program risk. By making reliability a requirement and asking suppliers to describe how they plan to achieve the required level of reliability and provide early predictions, suppliers have a basis for preparing their proposals. The customer, in turn, has a basis for evaluating each proposal for the level of risk, and in selecting the "best value" supplier. Of course, reliability is just one consideration in source selection.

f. *Identify and rank potential problem areas and suggest possible solutions.* In the course of design and development test, many problems will emerge. Some of these will be critical and the program cannot proceed until and unless they are solved. Many others, however, will not fall into this "critical" category. With limited time and resources, the issue is to prioritize these problems. Using predictions to determine which problems contribute most to unreliability facilitates the prioritization process.

g. *Provide a basis for selecting economic warranty period.* For many products, warranty is an important subject. Although most commonly associated with commercial products, some military systems and equipment is procured with a warranty. The cost of the warranty is included in the price of the product of system. The question that the supplier must address is how much to charge for the warranty and for how long a period to warrant the product. Predicting the reliability is an important method for projecting the number of returns or claims under the warranty (using past experience is another method). Based on the number of projected claims, and the reliability as a function of time, the optimum warranty period, as well as the price, of the warranty can be determined.

h. *Determine spares requirements.* Whether it is one's personal automobile or the power generation system in a C4ISR facility, failures will occur. The failed items must be repaired or replaced. The latter requires spare parts or assemblies. In addition, some items will be replaced on a regular basis, as part of a preventive maintenance program. Again, spares are needed. Predictions play an important role in determining how many spares of each type are needed.

3-3. The basics

When designing a new product or system, it is difficult, impractical, and sometimes impossible to predict the reliability of the entire product in one step. It is more common to predict the reliability of individual subsystems, assemblies, or even parts and then to "sum" up the individual reliabilities to assess the overall product reliability. It is very much like estimating the weight of a product. One would first estimate (or perhaps know from past experience or from supplier specifications) the weights of all the individual items that make up the product. By summing them up, the weight of the product can be estimated. Of course, as we will see, the process of "summing" individual reliabilities is more complicated than simply adding the reliabilities together.

a. *Hazard function*. The probability that an item will fail in the next instant of time, given that it has not yet failed, is called the hazard function, which is the probability of failure as a function of time. For parts that wear out, gears for example, the hazard function increases with time. That is, the probability of failure is continuously increasing with time. For many items that do not wear out, the hazard function is constant with time. A system under development, for which design improvements are being made as a result of failures found during test or analysis, will have a decreasing hazard function. A system that is used beyond its designed useful life will begin to exhibit an increasing hazard function.

b. *Failure rate.* If the hazard function is constant, the probability of failure is constant over time. In such cases, it is commonly to use the term "failure rate" instead of hazard function. The hazard function is constant when the times to failure follow the exponential probability density function (pdf). It is also true that systems tend to behave as if the times to failure are exponentially distributed even if some parts within the system do not (i.e., they wear out). The reason is that systems are made up of may different types of parts, each type having its own underlying pdf for times to failure. As a result, the system behaves as if it has a failure rate, the inverse of which if the mean time between failure (MTBF). Of course, this is true only if the system is not under development (decreasing hazard function) or being used beyond its useful life (increasing hazard function).

c. *Basic reliability versus mission reliability prediction.* Many parts or assemblies in a system do not affect the system's ability to perform one or more of its functions. For example, the loss of one pump will not affect fluid flow if there is another pump that can take over. Even though the mission can be performed, the failed pump must be repaired (or replaced). Otherwise, another failure (of the other pump) will result in a mission failure. When we are interested in all failures, to determine spares and maintenance labor requirements, for example, we are addressing basic reliability, also called logistics reliability. When we are interested in only those failures that cause the mission to fail, we are addressing mission reliability. This distinction is important for many reasons. One of these, as we will see, is that the methods used for increasing mission reliability can actually cause basic reliability to decrease.

d. *Prediction iteration*. Reliability prediction for an individual component or an entire system is a process. Just as the design of a system evolves, with the designer going from a functional requirement to the physical solution, so the reliability prediction must evolve. Initially, data may not be available and predictions methods are based on similarity or generic part failure rates. As data becomes available, methods that capitalize on the data should be used. During design, this data will be the type, quantity, and quality of parts used, the manner in which the parts interface, the method of assembly and production, and the operational environment. As prototype/test products are available, actual operation/failure information can be gained from testing. Each iteration of the reliability prediction builds on previous work, adding the benefit of current information. The original estimate is very general, based on broad observations and is, therefore, itself very general. Each subsequent prediction, however, is based on more specific information, builds on the previous information, and the amount of uncertainty associated with the prediction decreases. After the demise of a system, the total failures, operating hours, etc., could be actually counted and the final and actual reliability calculated. In a very real sense then, we can visualize the prediction process as progressing from very crude estimates to an exact number. Seldom, however, can we extract every single bit of required data for even a retired system. Even when it is possible, such an exact number only serves as the broad basis to predict the reliability of a new, similar system. During the development and acquisition of a new system, we must recognize the uncertainty associated with any estimate.

3-4. Prediction method

A prediction can be made using a variety of methods, each with its own set of constraints and advantages. No one method is applicable for a product throughout its life cycle. A discussion of some of the most widely used and accepted methods follows. Examples of methods a. through e. are given in appendix C.

a. *Parts count.* This method uses the failure rates of the individual elements in a higher-level assembly to calculate the assembly failure rate. Note that in using failure rates, we are implicitly assuming that the times to failure are exponentially distributed. In using the parts count method, it is important that all portions of the higher-level assembly are used in and exposed to the same environment. The failure rates used can be based on first-hand experience and observation but are often the rates for generic part types. These generic failure rates are available from various sources, such as the Reliability Analysis Center, for a wide range of electronic and mechanical parts. As discussed in Paragraph 3-4e, these rates often are a cumulative average and the actual hazard function is not constant.

b. *Similarity analysis.* If a new product or system is being developed that is similar to a current product or system, the reliability prediction for the new product or system can be based on the current one. Of course, some "adjustments" must be made to account for any differences in the technology being used, the way in which the product will be used, and any differences in the operating environment. Although such adjustments are not exact, similarity analysis is a good way to obtain a very early estimate of the level of reliability that can be expected. Even if the entire product is not similar to an existing one, subsystems or assemblies may be. Often, a specific pump, generator, or other component will be used in different systems. If the operating environment and usage is similar, then the reliability of the component in one system should be similar to the reliability in another system.

c. *Stress-strength interference method*. This method can be used to obtain a point estimate of reliability for an unlimited number of mechanical components. Stress and strength are treated as random variables defined by probability density functions (pdfs). As shown in figure 3-1, the curves for the two pdfs overlap forming an area of interference. (Note that although the curves shown in the figure are of two Normal distributions, the actual pdfs for stress and strength can be any distribution.) The interference is equal to the unreliability (i.e., a weak part meeting a high stress).

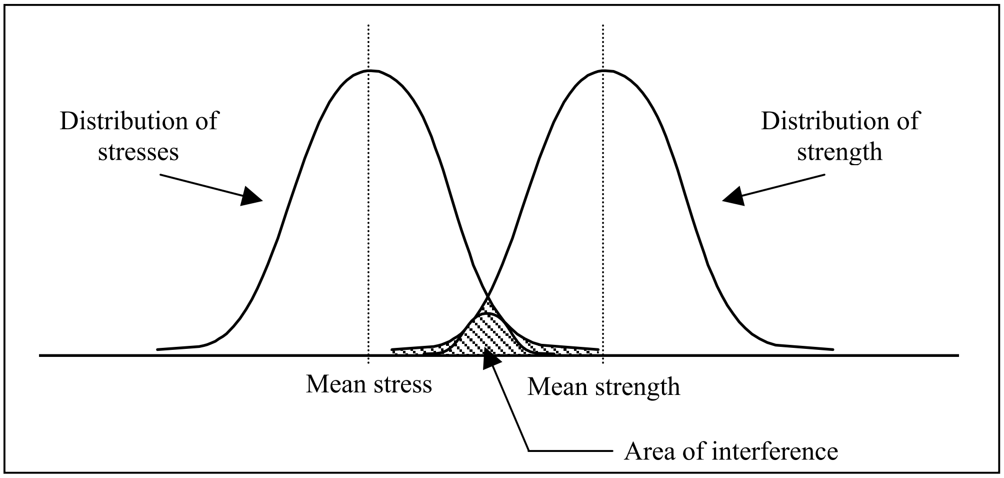


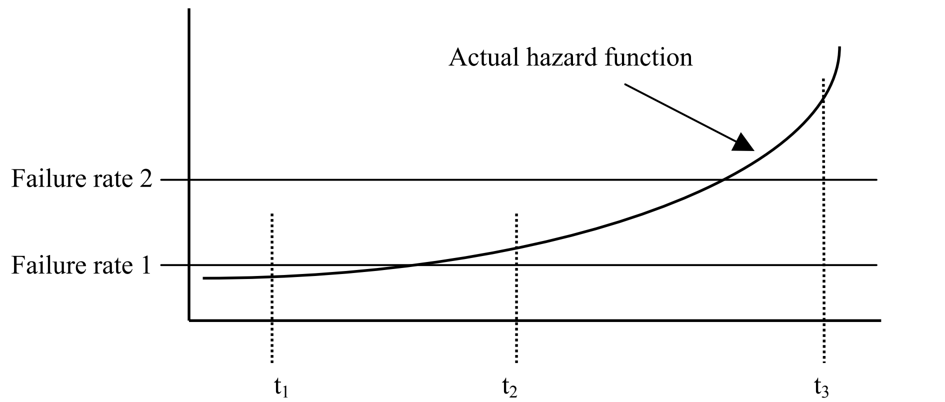
Figure 3-1. The area of interference in the stress-strength interference method is the probability of failure (the unreliability).

d. *Empirical models*. Models and formulas are available for many components that are based on actual data observed over a range of conditions. These models are sensitive to and only valid for the primary variables causing failure. A point estimate of reliability can be obtained at points of interest, allowing design trade-offs to be made early in the design phase. Table 3-1 describes two of the more common empirical models used today.

Table 3-1. Two empirical models for predicting reliability

Model Type	Equation or Model	Notes
Bearing Life Prediction	$B_{10} = \left(\frac{C}{P}\right)^{K} \text{ x } 10^{6}$ revolutions	B_{10} is the number of revolutions at which 90% of a population of bearings would survive. C is the load rating and K is a factor that varies depending on the type of bearing. C and K come from the manufacturer's literature.
Fatigue Curves	Curves that indicate fatigue life of a material in number of stress cycles before failure.	Curves are available for many ferrous and non-ferrous alloys, can reflect the effect of surface hardening, crack growth rate, effects of environmental stress variables, stress risers (e.g., holes), etc.

e. *Failure data analysis*. When data are available from test or form field use, the data can be used to assess the reliability of the item. When the data are for part failures, a valve for example, and the times to each failure have been collected, Weibull analysis can be used. The Weibull is a probability density function developed by a Swedish engineer Waloddi Weibull, who was studying fatigue failures. Weibull analysis is a powerful tool that can be used when the underlying distribution of the times to failure are actually Weibull, Normal, or exponential. It can be used when a lot of test or operating time has been accumulated but very few failures have been observed. Often, the times to failure are not known. In this case, we will know only the total time accumulated and the total number of failures. This type of data is called grouped data. Using grouped data, an average failure rate, the total number of failures divided by the total time, can be used. This rate actually represents a cumulative average that is valid for the time period over which the data were collected. If the hazard function for the part is actually increasing, the cumulative average will change depending on the period of interest. Figure 3-2 illustrates how grouped data is used to calculate a cumulative average failure rate.



Failure rate 1 is the average over t_1 to t_2. Failure rate 2 is the average over t_2 to t_3.

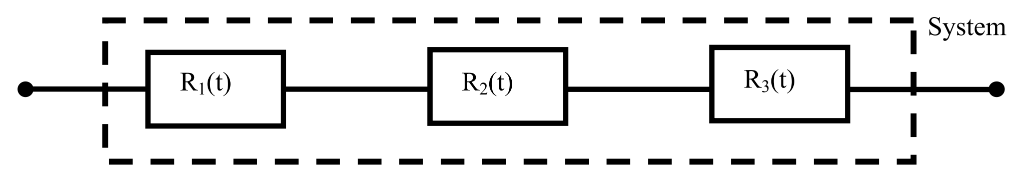
Figure 3-2. The relationship of the average cumulative failure rate and the actual hazard function for a part.

f. *Trending*. When monitoring the reliability of systems under development or in use, it is useful to determine if the system reliability is staying the same, getting worse, or improving. During development, as the design matures, one would expect the reliability be improving. As a system approaches the end of its useful life, one would expect the system reliability to start degrading. Between the end of design and the end of the useful life, one would expect the reliability to stay the same. It will stay constant unless some change is made. The change could be a change in the way the system is used, a change in a manufacturing process or source of a part or assembly, or a change in the competency level of the operators or maintainers. Many techniques exist for performing trending. One of these will be discussed in chapter 6.

3-5. Reliability modeling

Parts and assemblies can be connected in several different configurations. A reliability model is a way of depicting the connections from a reliability perspective. The most common modeling approach used today is the Reliability Block Diagram (RBD). The RBD consists of three basic types of building blocks: series configurations, parallel configurations, and combinations of series and parallel configurations.

a. *Series configuration.* The simplest way to think of a series configuration is as a chain. Just as a chain is only as strong as its weakest link, so the reliability of a series configuration is limited by the least reliable element in the series. For example, if a road crosses three bridges, the loss of any one bridge will prevent traffic from moving. Figure 3-3 shows a simple series configuration and how the system reliability is calculated using the reliability of each element.

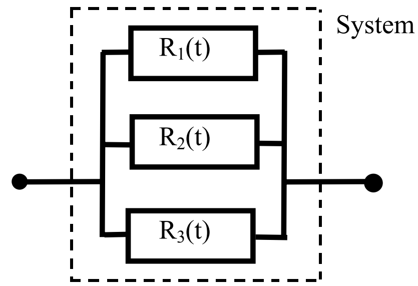


$$R_{System}(t) = R_1(t) \times R_2(t) \times R_3(t)$$

Figure 3-3. The reliability of a system when all the elements in the system are in series is the product of the individual reliabilities.

b. *Parallel (or redundant) configuration.* In a parallel configuration, two or more alternate paths are available for performing a function. Consider the following example. If a road comes to a river that has three bridges over it, traffic can cross over any of the bridges, and any one bridge is sufficient to carry the amount of traffic that crosses each day, then all three bridges would have to fail before traffic would stop. The three bridges are said to be in parallel configuration, and this configuration is obviously more reliable than a series configuration, in which the failure of only one bridge will cause the flow of traffic to stop. Many types of parallel configurations can be used. Brief descriptions of three of these configurations follow.

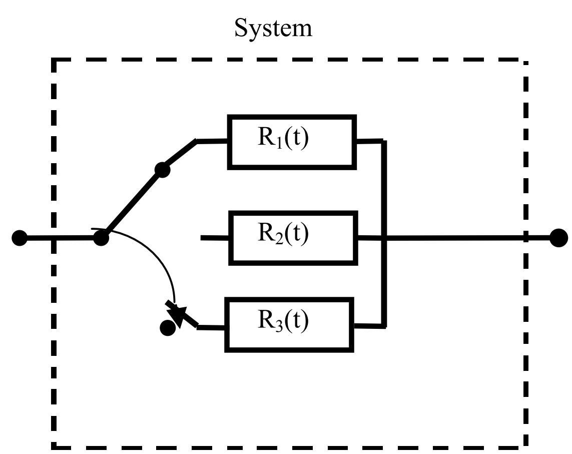
(1) Active parallel configuration (redundancy): all elements are on all of the time that the system is on and are immediately available to take over the function in the event any one element fails. The easiest way to calculate the reliability of the configuration is to determine the probability of all failing, and then to subtract this probability from 1. See figure 3-4.



$$R_{System}(t) = 1- ([1-R_1(t)] \times [1-R_2(t)] \times [1-R_3(t)])$$

Figure 3-4. In an active parallel configuration, the system reliability is calculated by multiplying the unreliability of the elements and subtracting the product from 1.

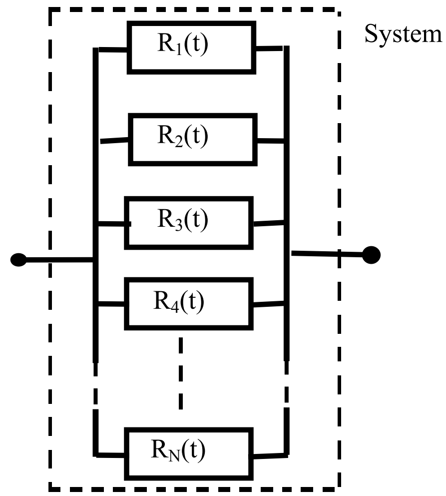
(2) Standby parallel configuration (redundancy): one element is performing the necessary function and another element must be switched on in the event of failure. In this configuration, there must be some method for detecting a failure and switching in the parallel element. Since the switch can fail, this configuration introduces additional opportunities for failure. The other element may be operating or not. If it is not, then the switching capability must also include some way of powering the inactive element on. Figure 3-5 shows this configuration with the reliability calculation when the switching is perfect (i.e., reliability of the switch is 100%), the standby elements are unpowered, and the times to failure for each of the elements are exponentially distributed (i.e., constant hazard function).



$$R_{System}(t) = e^{-\lambda t} + \lambda t e^{-\lambda t} + .5(\lambda t)^2 e^{-\lambda t}$$

Figure 3-5. Calculating the reliability of a parallel configuration with perfect switching, unpowered standby elements, and constant hazard function for each parallel element.

(3) k of N parallel configuration (redundancy): several elements are in parallel and two or more (but less than all) of the elements are needed to perform the function. See figure 3-6.

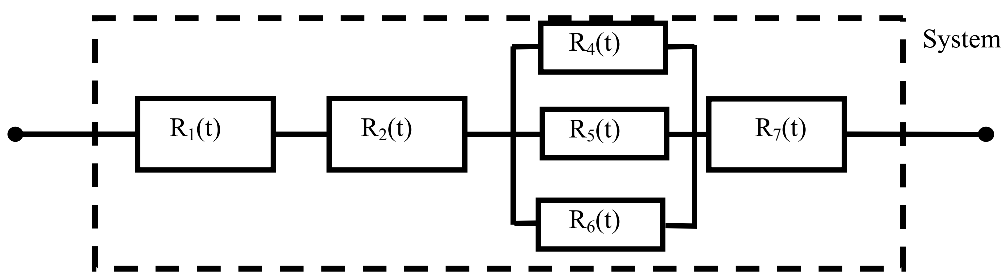


If $R_1(t) = R_2(t) = R_3(t) = R_4(t) = \bullet \bullet \bullet = R_N(t) = R(t)$

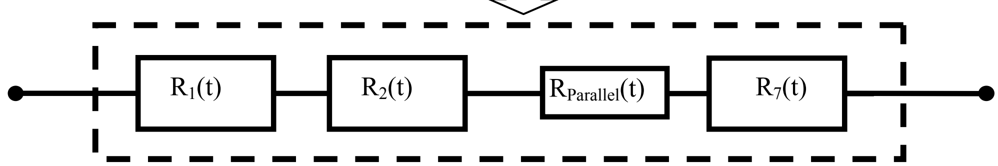
$$R_{System} = \sum_{i=K}^{N} \frac{N!}{i!\,(N-i)!} \left(R\right)^{i} \left(1 - R\right)^{N-i}$$

Figure 3-6. Calculating the reliability of k of N parallel elements of equal reliability.

c. *Combined configuration*. Any combination of series and the various parallel configurations is possible. To calculate the system reliability, first calculate the reliability of each individual configuration. The result is a series configuration for which the reliabilities can be multiplied to find the system reliability. See figure 3-7 for a simple example of a combined configuration.



$R_{Parallel}(t) = 1- ([1-R_4(t)] \times [1-R_5(t)] \times [1-R_6(t)])$



$R_{System}(t) = R_1(t) \times R_2(t) \times R_{Parallel}(t) \times R_7(t)$

Figure 3-7. Calculating the reliability of a combined configuration.